The Grid Analysis Environment
WorkPlan in Association with the PPDG Proposal
PPDG-38, Jan 2004

As part of the LCG ARDA project [ARDA], CMS will continue to work toward delivering an ARDA prototype and define a plan for deploying, developing and integrating ARDA services within the context of the Grid Analysis Environment work. Within ARDA, the Grid Analysis Environment [GAE] work has been identified by CMS, as vital in creating a distributed analysis environment. Requirements for interfaces between Grid services and CMS metadata services, CMS data model for distributed analysis will be defined within a comprehensive work-plan.

The Grid-Enabled Analysis Environment (GAE), in its final form will represent an end to end system for scalable distributed multi user (batch and interactive) analysis, in which Clarens web service hosts [CLAR] will be the backbone, on which services are deployed. Due to the distributed dynamic environment that GAE represents, the (web) server has to become a more active participant in this environment. Specifically seeking out server peers through lookup services and in turn advertising those services, such that users (and services) can dynamically discover services and resources within the GAE environment [HOS], This also implies more active internal state management of sessions, connections, as well as policy enforcement on resource usage. In tandem with the base functionality outlined above, specific services will be developed as part of the GAE, as well as the client applications that make use of these services. Specific emphasis will be given to interacting with the ROOT analysis framework [ROOT] , the CMS COBRA/ORCA framework [ORCA], and the POOL persistency framework [POOL]. GAE will use Chimera (VDT) for derived data management (section 5.5) Releases of server and client packages will be made as new functionality is implemented for testing and deployment as part of common software distributions such as the Virtual Data Toolkit (VDT), and supported in collaboration with the Clarens user community.

GAE will serve as a platform in which interfaces developed and discussed within the ppdg cs11 group can be tested, improve and evolve and will also interact with the ARDA workgroup on common components. Success of GAE will depend on early adapters of GAE for physics analysis. It is therefore important to work closely with physicists doing interactive or batch analysis in early stages of GAE for testing and feedback. Development of the Grid Analysis Environment at Caltech will continue in collaboration with the University of Florida and other U.S. CMS institutions.

Part of GAE work will be extensions to the Clarens Web Services Framework which acts as the GAE backbone. Clarens provides a specification for an implementation of Grid-enabled web services using standard protocols and tools. The Clarens framework is gaining acceptance and use by a variety of applications for LHC experiments. Development of the two implementations of Clarens in an open an collaborative fashion in response to the needs of the user community will continue, with a focus on interoperability, standards compliance and robustness. End-to-End Scalable Monitoring System. Based on MonaLisa services [MONA] will be added to interact and marshall data from diverse services for Grid monitoring [*MDS1], cluster monitoring [*Ganglia1, *GEMS1], network monitoring [*I2E2EPi], and to monitor and track the overall Grid system in real-time and the individual jobs submitted by users. The following milestones and deliverables have been identified within the context of GAE and PPDG:

Milestone 1: Set up existing components (Scheduler,Chimera VDT, Monalisa, replica location service), create a stable and easy to install system.. Specification of Lookup interface. Documentation of interfaces of existing components and discussion with cs11 group on more generic interfaces for these components. (April 2004).

Milestone 2: Enable clients to access the Pool meta data catalog via the Clarens web service environment. Enasble IGUANA (CMS visualization client) to browse meta-data in POOL catalog via web services. (May 2004) .

Milestone 3: First version of a browser oriented graphical user interface (workflow editor, analysis script submission, clarens management, service display) which will provide a natural interface to GAE services

and to enable early feedback on GAE functionality from the end user. Deployment of first version of Lookup service for dynamic discovery. Multiple instances of available services can be deployed. Users and services can discover services. Minimizing single point of failure. (July 2004). July 2004 is also the time set by the ARDA workgroup to evaluate the first prototypes for distributed analysis.

Milestone 4: Deploy McRunjob as a web service and a first version of workflow management. McRunJob is able to use ORCA within the GAE environment. Enable users to simultaneously perform interactive analysis and production/batch analysis. Multiple users will do interactive analysis while a smaller number of user will do batch/production analysis using root and/or ORCA. Results of the analysis are registered in the meta data catalog and chimera. Interactive analysis users will get more priority than batch analysis. (August 2004).

Milestone 5: Distributed scheduler (the scheduler itself is distributed) and replica location service. Deployment of second version of lookup service for dynamic discovery. Multiple instances of service will be deployed. Users and services can discover services. Minimizing single point of failure.(September 2004).

Milestone 6: Integrate first implementation of accounting and authorization as web service within GAE (see section 4.1.3 and 5.1.4) (January 2005).

Milestone 7: Enable user feedback on the status of their jobs within GAE (see section 5.1.5). Further integration of MonaLisa to support user feedback (July 2005).

Milestone 8: Second version of a browser oriented graphical user interface (integrated monitor feedback, application control via steering service). Second version of workflow component (based on MCRunJob) The GAE user interface environment will be able to interact with a McRunJob (Clarens) web service that will be used to provide initial workflow builder/management capabilities for (interactive) analysis (section 5.5). (July 2005).

Milestone 9: First implementation and deployment of distributed system optimization components embedded in existing GAE web services (e.g. Scheduling, Replica management) Multiple users perform different types of analysis, in time the pattern (type of analysis, and data to be analyzed) will change, Schedulers, Replication management and optimization services adjust to cater for these changed analysis patterns. (August 2005).

Milestone 10: Further intergration of trouble shooting functionality (see section 5.1.5) (June 2006).

Milestone 11: Integrate second implementation of accounting and authorization as web service within GAE (see section 4.1.3 and 5.1.4) Second implementation and deployment of distributed system optimization components embedded in existing GAE web services (e.g. Scheduling, Replica management) once first-round effective strategies are developed, self-learning adaptive optimization algorithms, such as self-organizing neural networks [SONN1] will be used to optimize system performance, according to metrics that take into account the policies for priority and resource usage of the Virtual Organization [NEW].
(December 2006).

Deployment and testing of the software will take place in 3 stages.

Milestone 1, Milestone 2: Deployed on the development grid testbed (DGT)
Milestone 2, Milestone 3, Milestone 4, Milestone 5: Integration Grid Testbed (IGT)
Milestone 5 until Milestone 11: Production Grid Testbed (PGT)

Users:
Milestone 1, Milestone 2: six to seven users that are doing analysis by submitting ROOT scripts
Milestone 3, until Milestone 5: fifteen to twenty users doing analysis either submitting Root scripts or using ORCA for either batch or interactive analysis.

Milestone 5 until Milestone 11: Initially starting with 20 users and ramping up to more than 100 user and beyond, doing a variety of analysis between pure batch and pure interactive using ROOT, ORCA and other analysis tools.

In cases where we can not get sufficient number of users for the milestones, we can have participating users submitting multiple analysis tasks to “mimic” multiple user.

Data:
???

References:

[ARDA] http://www.uscms.org/s&c/lcg/ARDA/, 2003
[GAE] web page, http://ultralight.caltech.edu/gaeweb/, 2003
[ClAR] C. Steenberg, E. Aslakson, J. Bunn, H. Newman, M. Thomas, F. van Lingen, "The Clarens Web Services Architecture", Proceedings of CHEP, La Jolla, California, March 2003
[HOS] W. Hoschek, A Unified Peer-to-Peer Database Framework for XQueries over Dynamic Distributed Content and its Application for Scalable Service Discovery (Dissertation), Technical University of Vienna, 2003
[ROOT] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework, Proceedings AIHENP'96 Workshop", Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86
[ORCA] "Object Oriented Reconstruction for CMS Analysis" cms internal note submitted to the LHCC referees in January 1999
[MONA] H.B. Newman, I.C. Legrand, P. Galvez, R. Voicu, C. Cirstioiu, "MonaALISA: A Distributed Monitoring Service Architecture", Proceedings of CHEP, La Jolla, California, March 2003.
[NEW] H. B. Newman, I. C. Legrand, "A Self-Organizing Neural Network for Job Scheduling in Distributed Systems", Presented at: ACAT 2000, FERMILAB, USA, October 16-20, 2000